

Modal Characterisations of Behavioral Pseudometrics

Yuxin Deng¹, Wenjie Du², and Daniel Gebler³

¹ East China Normal University

² Shanghai Normal University

³ VU University Amsterdam

14/09/2015

Abstract. For the model of probabilistic labelled transition systems that allow for the co-existence of nondeterminism and probabilities, we present two notions of bisimulation metrics: one is state-based and the other is distribution-based. We provide a sound and complete modal characterisation for each of them, using real-valued modal logics based on the Hennessy-Milner logic. The logic for characterising the state-based metric is much simpler than an earlier logic by Desharnais et al. as it uses only two non-expansive operators rather than the general class of non-expansive operators.

Keywords: Probabilistic labelled transition systems, behavioral pseudometrics, real-valued modal logics

1 Introduction

Bisimulation is an important proof technique for establishing behavioural equivalences of concurrent systems. In probabilistic concurrency theory, there are roughly two kinds of bisimulations: one is state-based because it is directly defined over states and then lifted to distributions, and the other is distribution-based as it is a relation between distributions. The former is originally defined in [LS91] to represent a branching time semantics; the latter, as defined in [HKK14, FZ14, DFD15], is strictly coarser and represents a linear time semantics.

In correspondence with those bisimulations, there are two notions of behavioural pseudometrics (simply called metrics). They are more robust ways of formalising behavioural similarity between formal systems than bisimulations because, particularly in the probabilistic setting, bisimulations are too sensitive to probabilities (a very small perturbation of the probabilities would render two systems non-bisimilar). A metric gives a quantitative measure of the distance between two systems and distance 0 usually means that the two systems are bisimilar. A logical characterisation of the state-based bisimulation metric for labelled Markov processes is given in [DGJP04]. For a more general model of labelled

concurrent Markov chains (LCMCs) that allow for the co-existence of nondeterminism and probabilities, a weak bisimulation metric is proposed in [DJGP02]. Its logical characterisation uses formulae like $h \circ f$, where f is a formula and h can be any non-expansive operator on $[0, 1]$, i.e. $|h(x) - h(y)| \leq |x - y|$ for any $x, y \in [0, 1]$. A natural question then arises: instead of the general class of non-expansive operators, is it possible to use only a few simple non-expansive operators without losing the capability of characterising the bisimulation metric?

In the current work, we give a positive answer to the above question. We work in the framework of probabilistic labelled transition systems (pLTSs) that are essentially the same as LCMCs, so the interplay of nondeterminism and probabilities is allowed. We provide a modal characterisation of the state-based bisimulation metric closely in line with the classical Hennessy-Milner logic (HML) [HM85]. Our variant of the HML makes use of state formulae and distribution formulae, which are formulae evaluated at states and distributions, respectively, and yield success probabilities. We use merely two non-expansive operators: negation ($\neg\phi$) and testing ($\phi \ominus p$). Negation is self-explanatory and the testing operator checks if a state satisfies a property with certain threshold probability. More precisely, if state s satisfies formula ϕ with probability q , then it satisfies $\neg\phi$ with probability $1 - q$, and satisfies $\phi \ominus p$ with probability $q - p$ if $q > p$ and 0 otherwise. In other words, we do not need the general class of non-expansive operators because negation and testing, together with other modalities in the classical HML, are expressive enough to characterise bisimulation metrics. As regards to the characterisation of distribution-based bisimulation metric, we drop state formulae and use distribution formulae only.

The rest of this paper is organised as follows. Section 2 provides some basic concepts on pLTSs. Section 3 defines a two-sorted modal logic that leads to a sound and complete characterisation of the state-based bisimulation metric. Section 4 gives a similar characterisation for the distribution-based bisimulation metric. In Section 5 we review some related work. Finally, we conclude in Section 6.

2 Preliminaries

Let S be a countable set. A (*discrete*) *probability subdistribution* over S is a function $\Delta : S \rightarrow [0, 1]$ with $\sum_{s \in S} \Delta(s) \leq 1$. It is a (*full*) *distribution* if $\sum_{s \in S} \Delta(s) = 1$. Let $\mathcal{D}(S)$ denote the set of all distributions over S . A matching $\omega \in \mathcal{D}(S \times S)$ for $(\Delta, \Theta) \in \mathcal{D}(S) \times \mathcal{D}(S)$ is given if $\sum_{t \in S} \omega(s, t) = \Delta(s)$ and $\sum_{s \in S} \omega(s, t) = \Theta(t)$ for all $s, t \in S$. We denote the set of all matchings for (Δ, Θ) by $\Omega(\Delta, \Theta)$.

A *metric* d over space S is a distance function $d : S \times S \rightarrow \mathbb{R}_{\geq 0}$ satisfying: (i) $d(s, t) = 0$ iff $s = t$ (isolation), (ii) $d(s, t) = d(t, s)$ (symmetry), (iii) $d(s, t) \leq d(s, u) + d(u, t)$ (triangle inequality), for any $s, t, u \in S$. If we replace (i) with $d(s, s) = 0$ for all $s \in S$, we obtain a *pseudometric*. In this paper we are interested in pseudometrics because two distinct states can still be at distance zero if their behaviour is similar. But for simplicity, we often use metrics for pseudometrics.

A metric d over S is c -bounded if $d(s, t) \leq c$ for any $s, t \in S$, where $c \in \mathbb{R}_{\geq 0}$ is a positive real number.

Let $d: S \times S \rightarrow [0, 1]$ be a metric over S . We lift it to be a metric over $\mathcal{D}(S)$ by using the *Kantorovich metric* [KR58] $K(d): \mathcal{D}(S) \times \mathcal{D}(S) \rightarrow [0, 1]$ defined via a linear programming problem as follows:

$$K(d)(\Delta, \Theta) = \min_{\omega \in \Omega(\Delta, \Theta)} \sum_{s, t \in S} d(s, t) \cdot \omega(s, t)$$

for $\Delta, \Theta \in \mathcal{D}(S)$. The dual of the above linear programming problem is the following

$$\max \sum_{s \in S} (\Delta(s) - \Theta(s))x_s, \text{ subject to } \begin{array}{l} 0 \leq x_s \leq 1 \\ \forall s, t \in S: x_s - x_t \leq d(s, t) \end{array}.$$

The duality theorem in linear programming guarantees that both problems have the same optimal value.

Let $\hat{d}: \mathcal{D}(S) \times \mathcal{D}(S) \rightarrow [0, 1]$ be a metric over $\mathcal{D}(S)$. We lift it to be a metric over the powerset of $\mathcal{D}(S)$, written $\mathcal{P}(\mathcal{D}(S))$, by using the *Hausdorff metric* $H(\hat{d}): \mathcal{P}(\mathcal{D}(S)) \times \mathcal{P}(\mathcal{D}(S)) \rightarrow [0, 1]$ given as follows

$$H(\hat{d})(\Pi_1, \Pi_2) = \max\left\{ \sup_{\Delta \in \Pi_1} \inf_{\Theta \in \Pi_2} \hat{d}(\Delta, \Theta), \sup_{\Theta \in \Pi_2} \inf_{\Delta \in \Pi_1} \hat{d}(\Theta, \Delta) \right\}$$

for all $\Pi_1, \Pi_2 \subseteq \mathcal{D}(S)$, whereby $\inf \emptyset = 1$ and $\sup \emptyset = 0$.

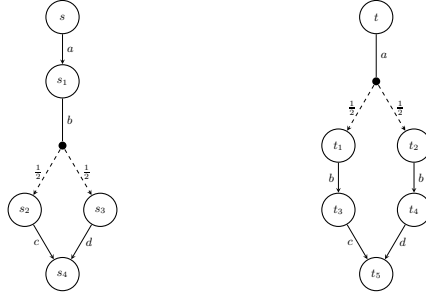
Probabilistic labelled transition systems (pLTSs) generalize labelled transition systems (LTSs) by allowing for probabilistic choices in the transitions. We consider pLTSs (or essentially *simple probabilistic automata* [Seg95]) with countable state spaces.

Definition 1. A probabilistic labelled transition system is a triple (S, A, \rightarrow) , where S is a countable set of states, A is a countable set of actions, and the relation $\rightarrow \subseteq S \times A \times \mathcal{D}(S)$ is a transition relation.

We write $s \xrightarrow{a} \Delta$ for $(s, a, \Delta) \in \rightarrow$ and define $der(s, a) = \{\Delta \mid s \xrightarrow{a} \Delta\}$ as the set of all a -successor distributions of s . A pLTS is *image-finite* if for any state s and action a the set $der(s, a)$ is finite. In the current work, we focus on image-finite pLTSs.

3 State-Based Bisimulation Metrics

We consider the complete lattice $([0, 1]^{S \times S}, \sqsubseteq)$ defined by $d \sqsubseteq d'$ iff $d(s, t) \leq d'(s, t)$, for all $s, t \in S$. For some $D \subseteq [0, 1]^{S \times S}$ the least upper bound is given by $(\bigsqcup D)(s, t) = \sup_{d \in D} d(s, t)$, and the greatest lower bound is given by $(\bigsqcap D)(s, t) = \inf_{d \in D} d(s, t)$ for all $s, t \in S$. The bottom element $\mathbf{0}$ is the constant zero function $\mathbf{0}(s, t) = 0$ and the top element $\mathbf{1}$ is the constant one function $\mathbf{1}(s, t) = 1$ for all $s, t \in S$.

Figure 1: $\mathbf{d}_{sb}(s, t) = \frac{1}{2}$

Definition 2. A 1-bounded metric d on S is a state-based bisimulation metric if for all $s, t \in S$ and $\epsilon \in [0, 1)$ with $d(s, t) \leq \epsilon$, if $s \xrightarrow{a} \Delta$ then there exists some $t \xrightarrow{a} \Delta'$ with $K(d)(\Delta, \Delta') \leq \epsilon$.

The smallest (wrt. \sqsubseteq) state-based bisimulation metric, denoted by \mathbf{d}_{sb} , is called *state-based bisimilarity metric*. Its kernel is the state-based bisimilarity as defined in [LS91, Seg95].

Example 3. In this example, we calculate the distance between states s and t in Figure 1. Firstly, observe that $\mathbf{d}_{sb}(s_2, t_3) = 0$ because s_2 is bisimilar to t_3 while $\mathbf{d}_{sb}(s_3, t_3) = 1$ because the two states s_3 and t_3 perform completely different actions. Secondly, let $\Delta = \frac{1}{2}\overline{s_2} + \frac{1}{2}\overline{s_3}$ and $\Theta = \overline{t_3}$. We see that

$$\begin{aligned} K(\mathbf{d}_{sb})(\Delta, \Theta) &= \min_{\omega \in \Omega(\Delta, \Theta)} \mathbf{d}_{sb}(s_2, t_3) \cdot \omega(s_2, t_3) + \mathbf{d}_{sb}(s_3, t_3) \cdot \omega(s_3, t_3) \\ &= \min_{\omega \in \Omega(\Delta, \Theta)} 1 \cdot \omega(s_2, t_3) + 0 \cdot \omega(s_3, t_3) \\ &= \min_{\omega \in \Omega(\Delta, \Theta)} 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

It follows that $\mathbf{d}_{sb}(s_1, t_1) = \frac{1}{2}$. Similarly, we get $\mathbf{d}_{sb}(s_1, t_2) = \frac{1}{2}$. Then it not difficult to see that

$$K(\mathbf{d}_{sb})(\overline{s_1}, \frac{1}{2}\overline{t_1} + \frac{1}{2}\overline{t_2}) = \mathbf{d}_{sb}(s_1, t_1) \cdot \frac{1}{2} + \mathbf{d}_{sb}(s_1, t_2) \cdot \frac{1}{2} = \frac{1}{2}$$

from which we finally obtain $\mathbf{d}_{sb}(s, t) = \frac{1}{2}$. \square

The above coinductively defined bisimilarity metric can be reformulated as a fixed point of a monotone functor. Let us define the functor $F: [0, 1]^{S \times S} \rightarrow [0, 1]^{S \times S}$ for $d: S \times S \rightarrow [0, 1]$ and $s, t \in S$ by

$$F(d)(s, t) = \sup_{a \in A} \{H(K(d))(der(s, a), der(t, a))\}.$$

It can be shown that F is monotone and its least fixed point is defined by $\bigcap d_i$, where $d_0 = \mathbf{0}$ and $d_{i+1} = F(d_i)$ for all $i \in \mathbb{N}$.

Proposition 4. *For image-finite pLTSs, the strong bisimilarity metric is the least fixed point of F .* \square

Now we proceed by defining a real-valued modal logic based on the Hennessy-Milner logic [HM85], called metric HML, to characterize the bisimilarity metric. Our logic is motivated by [DJGP02, Def. 4.1], [DGJP04, Def. 4.1] and [DTW10, Sec. 4]. In the remainder of this section we confine ourselves to pLTSs with finitely many states.

Definition 5. *Our metric HML is two-sorted and has the following syntax:*

$$\begin{aligned}\varphi &::= \top \mid \neg\varphi \mid \varphi \ominus p \mid \varphi_1 \wedge \varphi_2 \mid \langle a \rangle \psi \\ \psi &::= \psi \ominus p \mid \psi_1 \wedge \psi_2 \mid [\varphi]\end{aligned}$$

with $a \in A$ and $p \in [0, 1]$.

Let \mathcal{L} denote the set of all metric HML formulae, φ range over the set of all *state formulae* \mathcal{L}^S , and ψ range over the set of all *distribution formulae* \mathcal{L}^D . The two kinds of formulae are defined simultaneously. The operator $\phi \ominus p$ tests if a state passes ϕ with probability at least p . If ϕ is a state formula then it immediately induces a distribution formula $[\phi]$. Sometimes we abbreviate $\langle a \rangle[\varphi]$ as $\langle a \rangle\varphi$. All other operators are standard and have appeared in the classical HML.

Definition 6. *A state formula $\varphi \in \mathcal{L}^S$ evaluates in $s \in S$ as follows:*

$$\begin{aligned}\llbracket \top \rrbracket(s) &= 1, \\ \llbracket \neg\varphi \rrbracket(s) &= 1 - \llbracket \varphi \rrbracket(s), \\ \llbracket \varphi \ominus p \rrbracket(s) &= \max(\llbracket \varphi \rrbracket(s) - p, 0), \\ \llbracket \varphi_1 \wedge \varphi_2 \rrbracket(s) &= \min(\llbracket \varphi_1 \rrbracket(s), \llbracket \varphi_2 \rrbracket(s)), \\ \llbracket \langle a \rangle \psi \rrbracket(s) &= \max_{s \xrightarrow{a} \Delta} \llbracket \psi \rrbracket(\Delta),\end{aligned}$$

and a distribution formula $\psi \in \mathcal{L}^D$ evaluates in $\Delta \in \mathcal{D}(S)$ as follows:

$$\begin{aligned}\llbracket \psi \ominus p \rrbracket(\Delta) &= \max(\llbracket \psi \rrbracket(\Delta) - p, 0), \\ \llbracket \psi_1 \wedge \psi_2 \rrbracket(\Delta) &= \min(\llbracket \psi_1 \rrbracket(\Delta), \llbracket \psi_2 \rrbracket(\Delta)), \\ \llbracket [\varphi] \rrbracket(\Delta) &= \sum_{s \in S} \Delta(s) \llbracket \varphi \rrbracket(s).\end{aligned}$$

We often use constant formulae e.g. \underline{p} for any $p \in [0, 1]$ with the semantics $\llbracket \underline{p} \rrbracket(s) = p$, which is derivable in the above logic by letting $\underline{p} = \top \ominus (1 - p)$. Moreover, we write $\varphi \oplus p$ for $\neg((\neg\varphi) \ominus p)$ which has the semantics $\llbracket \varphi \oplus p \rrbracket(s) = \min(\llbracket \varphi \rrbracket(s) + p, 1) = 1 - \max(1 - \llbracket \varphi \rrbracket(s) - p, 0)$. In the presence of negation and conjunction we can derive disjunction by letting $\varphi_1 \vee \varphi_2$ to be $\neg(\neg\varphi_1 \wedge \neg\varphi_2)$.

Semantics of $\langle a \rangle \varphi$ is a translation of [DJGP02, Def. 4.1] from labelled concurrent Markov chain semantics to pLTSs. Conjunction of distribution formulae $\psi_1 \wedge \psi_2$ could alternatively be replaced by $\psi_1 \oplus_p \psi_2$ or $\psi_1 \oplus \psi_2$ [Hen12, Sec. 4] with semantics

$$\begin{aligned} \llbracket \psi_1 \oplus_p \psi_2 \rrbracket(\Delta) &= \sup_{\Delta = p\Delta_1 + (1-p)\Delta_2} (p \cdot \llbracket \psi_1 \rrbracket(\Delta_1) + (1-p) \llbracket \psi_2 \rrbracket(\Delta_2)) \\ \llbracket \psi_1 \oplus \psi_2 \rrbracket(\Delta) &= \sup_{p \in (0,1)} \llbracket \psi_1 \oplus_p \psi_2 \rrbracket(\Delta) \end{aligned}$$

The above metric HML induces two natural logical metrics \mathbf{d}_{sb}^{ls} and \mathbf{d}_{sb}^{ld} on states and distributions respectively, by letting

$$\begin{aligned} \mathbf{d}_{sb}^{ls}(s, t) &= \sup_{\varphi \in \mathcal{L}^S} |\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| \\ \mathbf{d}_{sb}^{ld}(\Delta, \Theta) &= \sup_{\psi \in \mathcal{L}^D} |\llbracket \psi \rrbracket(\Delta) - \llbracket \psi \rrbracket(\Theta)| \end{aligned}$$

Example 7. Consider the two probabilistic systems depicted in Figure 2. We have the formula $\varphi = \langle a \rangle \psi$ where $\psi = [\langle a \rangle \top] \wedge [\langle b \rangle \top]$ and would like to know the difference s and t given by φ . Let

$$\begin{aligned} \Delta_1 &= 0.2 \cdot \overline{s_1} + 0.8 \cdot \overline{s_2} \\ \Delta_2 &= 0.8 \cdot \overline{s_5} + 0.2 \cdot \overline{s_6} \\ \Delta_3 &= 0.5 \cdot \overline{s_3} + 0.5 \cdot \overline{s_4} \end{aligned}$$

Note that $\llbracket \langle a \rangle \top \rrbracket(s_1) = 1$ and $\llbracket \langle a \rangle \top \rrbracket(s_2) = 0$. Then $\llbracket [\langle a \rangle \top] \rrbracket(\Delta_1) = 0.2 \cdot \llbracket \langle a \rangle \top \rrbracket(s_1) + 0.8 \cdot \llbracket \langle a \rangle \top \rrbracket(s_2) = 0.2$. Similarly, $\llbracket [\langle b \rangle \top] \rrbracket(\Delta_1) = 0.8$. It follows that $\llbracket \psi \rrbracket(\Delta_1) = \min(\llbracket [\langle a \rangle \top] \rrbracket(\Delta_1), \llbracket [\langle b \rangle \top] \rrbracket(\Delta_1)) = 0.2$. With similar arguments, we see that $\llbracket \psi \rrbracket(\Delta_2) = 0.2$ and $\llbracket \psi \rrbracket(\Delta_3) = 0.5$. Therefore, we can calculate that

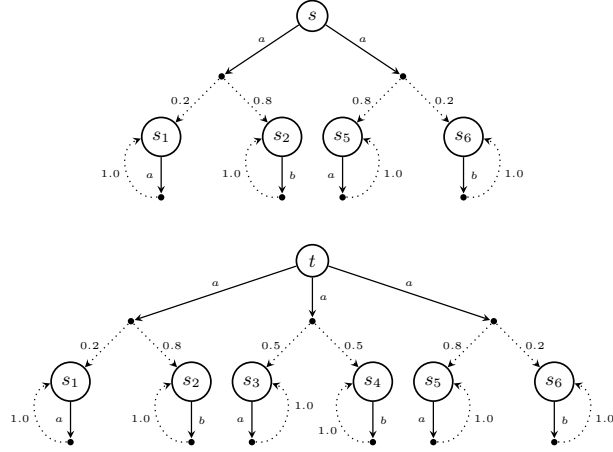
$$\begin{aligned} \llbracket \varphi \rrbracket(s) &= \max(\llbracket \psi \rrbracket(\Delta_1), \llbracket \psi \rrbracket(\Delta_2)) = 0.2 \\ \llbracket \varphi \rrbracket(t) &= \max(\llbracket \psi \rrbracket(\Delta_1), \llbracket \psi \rrbracket(\Delta_2), \llbracket \psi \rrbracket(\Delta_3)) = 0.5. \end{aligned}$$

So the difference between s and t with respect to φ is $|\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| = 0.3$. In fact we also have $\mathbf{d}_{sb}^{ls}(s, t) = 0.3$. □

Example 8. At first sight the following two equations seem to be sound.

$$\llbracket [\varphi] \ominus p \rrbracket(\Delta) = \llbracket [\varphi \ominus p] \rrbracket(\Delta) \quad \text{and} \quad \llbracket \psi \rrbracket\left(\sum_i p_i \Delta_i\right) = \sum_i p_i (\llbracket \psi \rrbracket(\Delta_i))$$

However, in general they do not hold, as witnessed by the counterexamples below. Let $\varphi = \langle b \rangle \top$, $\psi = [\varphi] \ominus 0.5$ and the distribution Δ_1 be the same as in

Figure 2: $\mathbf{d}_{sb}^{\text{ls}}(s, t) = 0.3$

Example 7. Then we have

$$\begin{aligned}
 \llbracket [\varphi] \ominus 0.5 \rrbracket(\Delta_1) &= \max(\llbracket [\varphi] \rrbracket(\Delta_1) - 0.5, 0) \\
 &= \max(0.2 \llbracket \langle b \rangle \top \rrbracket(\overline{s_1}) + 0.8 \llbracket \langle b \rangle \top \rrbracket(\overline{s_2}) - 0.5, 0) \\
 &= \max(0.2 \cdot 0 + 0.8 \cdot 1 - 0.5, 0) \\
 &= 0.3
 \end{aligned}$$

$$\begin{aligned}
 \llbracket [\varphi \ominus 0.5] \rrbracket(\Delta_1) &= 0.2 \llbracket [\varphi \ominus 0.5] \rrbracket(s_1) + 0.8 \llbracket [\varphi \ominus 0.5] \rrbracket(s_2) \\
 &= 0.2 \max(\llbracket [\varphi] \rrbracket(s_1) - 0.5, 0) + 0.8 \max(\llbracket [\varphi] \rrbracket(s_2) - 0.5, 0) \\
 &= 0.2 \max(0 - 0.5, 0) + 0.8 \max(1 - 0.5, 0) \\
 &= 0.4
 \end{aligned}$$

$$\begin{aligned}
 0.2 \llbracket [\psi] \rrbracket(\overline{s_1}) + 0.8 \llbracket [\psi] \rrbracket(\overline{s_2}) &= 0.2 \llbracket [\varphi] \ominus 0.5 \rrbracket(\overline{s_1}) + 0.8 \llbracket [\varphi] \ominus 0.5 \rrbracket(\overline{s_2}) \\
 &= 0.2 \max(\llbracket [\varphi] \rrbracket(\overline{s_1}) - 0.5, 0) + 0.8 \max(\llbracket [\varphi] \rrbracket(\overline{s_2}) - 0.5, 0) \\
 &= 0.2 \max(0 - 0.5, 0) + 0.8 \max(1 - 0.5, 0) \\
 &= 0.4
 \end{aligned}$$

So we see that $\llbracket [\varphi] \ominus 0.5 \rrbracket(\Delta_1) \neq \llbracket [\varphi \ominus 0.5] \rrbracket(\Delta_1)$ and $\llbracket [\psi] \rrbracket(\Delta_1) \neq 0.2 \llbracket [\psi] \rrbracket(\overline{s_1}) + 0.8 \llbracket [\psi] \rrbracket(\overline{s_2})$ \square

In what follows we will show that the logic \mathcal{L} precisely captures the bisimilarity metric \mathbf{d}_{sb} : the metric $\mathbf{d}_{sb}^{\text{ls}}$ defined by state formulae coincides with \mathbf{d}_{sb} and the metric $\mathbf{d}_{sb}^{\text{ld}}$ defined by distribution formulae coincides with $K(\mathbf{d}_{sb})$, the lifted form of \mathbf{d}_{sb} . The two properties are entangled because state formulae and distribution formulae are not independent.

Lemma 9. 1. $\mathbf{d}_{sb}^{\text{ls}} \subseteq \mathbf{d}_{sb}$
 2. $\mathbf{d}_{sb}^{\text{ld}} \subseteq K(\mathbf{d}_{sb})$

Proof. We show the two statements simultaneously by structural induction on formulae. For any two states $s, t \in S$ and distributions $\Delta_1, \Delta_2 \in \mathcal{D}(S)$, we prove that

- (1) $|\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| \leq \mathbf{d}_{sb}(s, t)$ for all $\varphi \in \mathcal{L}^S$;
- (2) $|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| \leq K(\mathbf{d}_{sb})(\Delta_1, \Delta_2)$ for all $\psi \in \mathcal{L}^D$.

We first analyze the structure of φ in (1).

- $\varphi \equiv \top$. Then it is trivial to see that $|\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| = |1 - 1| = 0 \leq \mathbf{d}_{sb}(s, t)$.
- $\varphi \equiv \neg \varphi'$. Then $|\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| = |\llbracket \varphi' \rrbracket(t) - \llbracket \varphi' \rrbracket(s)| \leq \mathbf{d}_{sb}(s, t)$ where the inequality holds by induction.
- $\varphi \equiv \varphi' \ominus p$. There are four subcases and we consider one of them. Suppose $\llbracket \varphi' \rrbracket(s) > p$ and $\llbracket \varphi' \rrbracket(t) \leq p$, then $|\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)| = |\llbracket \varphi' \rrbracket(s) - p| \leq |\llbracket \varphi' \rrbracket(s) - \llbracket \varphi' \rrbracket(t)| \leq \mathbf{d}_{sb}(s, t)$ by induction.
- $\varphi \equiv \varphi_1 \wedge \varphi_2$. Without loss of generality we assume that $\llbracket \varphi \rrbracket(s) \geq \llbracket \varphi \rrbracket(t)$. There are two possibilities:
 - If $\llbracket \varphi_1 \rrbracket(t) \leq \llbracket \varphi_2 \rrbracket(t)$, then $\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t) \leq \llbracket \varphi_1 \rrbracket(s) - \llbracket \varphi_1 \rrbracket(t) \leq \mathbf{d}_{sb}(s, t)$, where the last inequality holds by induction.
 - Symmetrically, if $\llbracket \varphi_2 \rrbracket(t) \leq \llbracket \varphi_1 \rrbracket(t)$, then $\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t) \leq \llbracket \varphi_2 \rrbracket(s) - \llbracket \varphi_2 \rrbracket(t) \leq \mathbf{d}_{sb}(s, t)$.
- $\varphi \equiv \langle a \rangle \psi$. We consider the non-trivial case that both s and t can perform action a . (If either of the two states cannot perform action a , the expected result is straightforward.) Let Δ_1 be the distribution such that $s \xrightarrow{a} \Delta_1$ and $\llbracket \langle a \rangle \psi \rrbracket(s) = \llbracket \psi \rrbracket(\Delta_1)$. Since \mathbf{d}_{sb} is a state-based bisimulation metric, by definition there exists some Δ_2 such that $t \xrightarrow{a} \Delta_2$ and $K(\mathbf{d}_{sb})(\Delta_1, \Delta_2) \leq \mathbf{d}_{sb}(s, t)$. Without loss of generality we assume that $\llbracket \varphi \rrbracket(s) \geq \llbracket \varphi \rrbracket(t)$. It follows that

$$\begin{aligned}
 & \llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t) \\
 &= \llbracket \psi \rrbracket(\Delta_1) - \max_{s \xrightarrow{a} \Delta'} \llbracket \psi \rrbracket(\Delta') \\
 &\leq \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \\
 &\leq K(\mathbf{d}_{sb})(\Delta_1, \Delta_2) \quad \text{by induction on } \psi \\
 &\leq \mathbf{d}_{sb}(s, t)
 \end{aligned}$$

Then we analyze the structure of ψ in (2).

- $\psi = \psi_1 \wedge \psi_2$. Without loss of generality we assume that $\llbracket \psi \rrbracket(\Delta_1) \geq \llbracket \psi \rrbracket(\Delta_2)$. There are two possibilities:
 - If $\llbracket \psi_1 \rrbracket(\Delta_2) \leq \llbracket \psi_2 \rrbracket(\Delta_2)$, then $\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \leq \llbracket \psi_1 \rrbracket(\Delta_1) - \llbracket \psi_1 \rrbracket(\Delta_2) \leq K(\mathbf{d}_{sb})(\Delta_1, \Delta_2)$, where the last inequality holds by induction.
 - Symmetrically, if $\llbracket \psi_2 \rrbracket(\Delta_2) \leq \llbracket \psi_1 \rrbracket(\Delta_2)$, then $\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \leq \llbracket \psi_2 \rrbracket(\Delta_1) - \llbracket \psi_2 \rrbracket(\Delta_2) \leq K(\mathbf{d}_{sb})(\Delta_1, \Delta_2)$.
- $\psi = \psi' \ominus p$ for some p in $[0, 1]$. There are four subcases and we consider one of them. Suppose $\llbracket \psi' \rrbracket(\Delta_1) > p$ and $\llbracket \psi' \rrbracket(\Delta_2) \leq p$, then $|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| = |\llbracket \psi' \rrbracket(\Delta_1) - p| \leq |\llbracket \psi' \rrbracket(\Delta_1) - \llbracket \psi' \rrbracket(\Delta_2)| \leq K(\mathbf{d}_{sb})(\Delta_1, \Delta_2)$ by induction.

- $\psi = [\varphi]$ for some $\varphi \in \mathcal{L}^S$. Without loss of generality we assume that $\llbracket \varphi \rrbracket(s) \geq \llbracket \varphi \rrbracket(t)$. We infer that

$$\begin{aligned}
& \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \\
&= \llbracket [\varphi] \rrbracket(\Delta_1) - \llbracket [\varphi] \rrbracket(\Delta_2) \\
&= \sum_{u \in S} (\Delta_1(u) - \Delta_2(u)) \llbracket \varphi \rrbracket(u) \\
&\leq \max \{ \sum_{u \in S} (\Delta_1(u) - \Delta_2(u)) x_u \mid x_u, x_{u'} \in [0, 1] \wedge x_u - x_{u'} \leq \mathbf{d}_{sb}(u, u') \} \\
&= K(\mathbf{d}_{sb})(\Delta_1, \Delta_2)
\end{aligned}$$

where the last equality holds because of the Kantorovich-Rubinstein duality theorem [KR58,vBW01a] and the last inequality holds because for any states $u, u' \in S$ we have $\llbracket \varphi \rrbracket(u), \llbracket \varphi \rrbracket(u') \in [0, 1]$ and $|\llbracket \varphi \rrbracket(u) - \llbracket \varphi \rrbracket(u')| \leq \mathbf{d}_{sb}(u, u')$ by induction on φ .

□

Lemma 10. $K(\mathbf{d}_{sb}^{\text{ls}}) \subseteq \mathbf{d}_{sb}^{\text{ld}}$

Proof. Let Δ_1, Δ_2 be any two distributions in \mathcal{L}^D . We show that

$$K(\mathbf{d}_{sb}^{\text{ls}})(\Delta_1, \Delta_2) \leq \sup_{\psi \in \mathcal{L}^D} |\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)|,$$

using an idea inspired by [DJGP02]. Let $L(\Delta_1, \Delta_2)$ be the optimal value of the following linear program

$$\begin{aligned}
& \max \sum_{s \in S} (\Delta_1(s) - \Delta_2(s)) x_s, \text{ subject to } 0 \leq x_s \leq 1 \\
& \forall s, t \in S: x_s - x_t \leq \mathbf{d}_{sb}^{\text{ls}}(s, t)
\end{aligned}$$

Suppose $\{k_s\}_{s \in S}$ be a set of real numbers that maximizes the above linear program to reach $L(\Delta_1, \Delta_2)$. Let $e = \min\{1 - k_t \mid k_t < 1 \text{ and } t \in S\}$ and $\epsilon > 0$ be any positive real number smaller than e . Hence, if $t \in S$ and $k_t < 1$ then

$$k_t + \epsilon < 1. \quad (1)$$

We construct some formula ψ such that

$$L(\Delta_1, \Delta_2) - \epsilon < \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2). \quad (2)$$

For any $s, t \in S$, we distinguish two cases:

1. If $k_s > k_t$, then $0 < k_s - k_t \leq \mathbf{d}_{sb}^{\text{ls}}(s, t)$. It is easy to see that there exists some formula φ_{st} such that

$$k_s - k_t < \llbracket \varphi_{st} \rrbracket(s) - \llbracket \varphi_{st} \rrbracket(t) + \epsilon. \quad (3)$$

or equivalently $\llbracket \varphi_{st} \rrbracket(t) - \llbracket \varphi_{st} \rrbracket(s) + k_s < k_t + \epsilon$. We define a new formula

$$\varphi'_{st} = \begin{cases} \varphi_{st} \ominus (\llbracket \varphi_{st} \rrbracket(s) - k_s) & \text{if } \llbracket \varphi_{st} \rrbracket(s) > k_s \\ \varphi_{st} \oplus (k_s - \llbracket \varphi_{st} \rrbracket(s)) & \text{otherwise.} \end{cases}$$

Let us compare φ'_{st} with k_t .

(a) If $\llbracket \varphi_{st} \rrbracket(s) > k_s$, then

$$\begin{aligned} \llbracket \varphi'_{st} \rrbracket(t) &= \max(\llbracket \varphi_{st} \rrbracket(t) - \llbracket \varphi_{st} \rrbracket(s) + k_s, 0) \\ &< \max(k_t + \epsilon, 0) \quad \text{by (3)} \\ &= k_t + \epsilon \end{aligned}$$

(b) Otherwise, $\llbracket \varphi'_{st} \rrbracket(t) = \min(\llbracket \varphi_{st} \rrbracket(t) + k_s - \llbracket \varphi_{st} \rrbracket(s), 1)$. By (3) we infer that $\llbracket \varphi_{st} \rrbracket(t) + k_s - \llbracket \varphi_{st} \rrbracket(s) < k_t + \epsilon$. Since $k_t < k_s \leq 1$ we infer by (1) that $k_t + \epsilon < 1$ and thus $\llbracket \varphi'_{st} \rrbracket(t) < k_t + \epsilon$.

In both (a) and (b) we have $\llbracket \varphi'_{st} \rrbracket(t) < k_t + \epsilon$, and it is also easy to see that $\llbracket \varphi'_{st} \rrbracket(s) = k_s$.

2. If $k_s \leq k_t$, then we simply set φ'_{st} to be the formula $\underline{k_s}$. As in the last case, we have $\llbracket \varphi'_{st} \rrbracket(s) = k_s$ and $\llbracket \varphi'_{st} \rrbracket(t) = k_s \leq k_t < k_t + \epsilon$.

In summary, the above reasoning says that for any $s, t \in S$ we can construct a formula φ'_{st} such that $\llbracket \varphi'_{st} \rrbracket(s) = k_s$ and $\llbracket \varphi'_{st} \rrbracket(t) < k_t + \epsilon$. Now let us define $\varphi'_s = \bigwedge_{t \in S} \varphi'_{st}$. It is easy to see that $\llbracket \varphi'_s \rrbracket(s) = k_s$ and $\llbracket \varphi'_s \rrbracket(t) < k_t + \epsilon$ for all $t \in S$. The latter implies $\max\{\llbracket \varphi'_s \rrbracket(t) \mid s, t \in S\} < k_t + \epsilon$. Then define $\varphi = \bigvee_{s \in S} \varphi'_s$. For all $t \in S$, we have

$$k_t = \llbracket \varphi'_t \rrbracket(t) \leq \llbracket \varphi \rrbracket(t) = \max\{\llbracket \varphi'_s \rrbracket(t) \mid s, t \in S\} < k_t + \epsilon.$$

Finally, we define $\psi = [\varphi]$. It follows that

$$\begin{aligned} \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) &= \llbracket [\varphi] \rrbracket(\Delta_1) - \llbracket [\varphi] \rrbracket(\Delta_2) \\ &= \sum_{t \in S} \Delta_1(t) \cdot \llbracket \varphi \rrbracket(t) - \sum_{t \in S} \Delta_2(t) \cdot \llbracket \varphi \rrbracket(t) \\ &\geq \sum_{t \in S} \Delta_1(t) \cdot k_t - \sum_{t \in S} \Delta_2(t) \cdot \llbracket \varphi \rrbracket(t) \\ &> \sum_{t \in S} \Delta_1(t) \cdot k_t - \sum_{t \in S} \Delta_2(t) \cdot (k_t + \epsilon) \\ &= \sum_{t \in S} (\Delta_1(t) - \Delta_2(t)) \cdot k_t - \sum_{t \in S} \Delta_2(t) \cdot \epsilon \\ &= L(\Delta_1, \Delta_2) - \epsilon \end{aligned}$$

as required. \square

The above property will be used to prove the following lemma.

Lemma 11. $\mathbf{d}_{sb} \sqsubseteq \mathbf{d}_{sb}^{\text{ls}}$

Proof. We show that $\mathbf{d}_{sb}^{\text{ls}}$ is a state-based bisimulation metric. Let s, t be any two states in S and ϵ be any real number in the interval $[0, 1)$ with $\mathbf{d}_{sb}^{\text{ls}}(s, t) \leq \epsilon$. Assume that $s \xrightarrow{a} \Delta_1$ is an arbitrarily chosen transition from s . Then state t must be able to perform action a too. Otherwise it is easy to see that $\mathbf{d}_{sb}^{\text{ls}}(s, t) = 1 > \epsilon$, which contradicts our assumption above. We need to show that there exists some transition $t \xrightarrow{a} \Delta_2$ with $K(\mathbf{d}_{sb}^{\text{ls}})(\Delta_1, \Delta_2) \leq \epsilon$. Suppose for a contradiction that no a -transition from t satisfies this condition. In other words, for each Δ_2^i with $t \xrightarrow{a} \Delta_2^i$ we have $K(\mathbf{d}_{sb}^{\text{ls}})(\Delta_1, \Delta_2^i) > \epsilon$. By Lemma 10, this means $\mathbf{d}_{sb}^{\text{ls}}(\Delta_1, \Delta_2^i) > \epsilon$. Then there must exist some formula $\psi_2^i \in \mathcal{L}^D$ such that $|\llbracket \psi_2^i \rrbracket(\Delta_1) - \llbracket \psi_2^i \rrbracket(\Delta_2^i)| > \epsilon$. Furthermore, we can strengthen this condition to the following one

$$\llbracket \psi_2^i \rrbracket(\Delta_1) - \llbracket \psi_2^i \rrbracket(\Delta_2^i) > \epsilon \quad (4)$$

because we can take the formula $\neg\psi_2^i$ in place of ψ_2^i in the case that $\llbracket\psi_2^i\rrbracket(\Delta_1) < \llbracket\psi_2^i\rrbracket(\Delta_2)$. Let $\varphi = \langle a \rangle \bigwedge_i (\psi_2^i \ominus \llbracket\psi_2^i\rrbracket(\Delta_2^i))$. We infer that

$$\begin{aligned} \llbracket\varphi\rrbracket(s) &= \max_{s \xrightarrow{a} \Delta} \llbracket\bigwedge_i \psi_2^i \ominus \llbracket\psi_2^i\rrbracket(\Delta_2^i)\rrbracket(\Delta) \\ &\geq \llbracket\bigwedge_i (\psi_2^i \ominus \llbracket\psi_2^i\rrbracket(\Delta_2^i))\rrbracket(\Delta_1) \\ &= \min_i \llbracket\psi_2^i \ominus \llbracket\psi_2^i\rrbracket(\Delta_2^i)\rrbracket(\Delta_1) \\ &= \llbracket\psi_2^k \ominus \llbracket\psi_2^k\rrbracket(\Delta_2^k)\rrbracket(\Delta_1) \quad \text{for some } k \\ &= \max(\llbracket\psi_2^k\rrbracket(\Delta_1) - \llbracket\psi_2^k\rrbracket(\Delta_2^k), 0) \\ &> \epsilon \quad \text{by (4)} \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \llbracket\varphi\rrbracket(t) &= \max_{t \xrightarrow{a} \Delta_2^i} \llbracket\bigwedge_j (\psi_2^j \ominus \llbracket\psi_2^j\rrbracket(\Delta_2^j))\rrbracket(\Delta_2^i) \\ &= \max_{t \xrightarrow{a} \Delta_2^i} \min_j \llbracket\psi_2^j \ominus \llbracket\psi_2^j\rrbracket(\Delta_2^j)\rrbracket(\Delta_2^i) \\ &= \max_{t \xrightarrow{a} \Delta_2^i} \min_j \max(\llbracket\psi_2^j\rrbracket(\Delta_2^i) - \llbracket\psi_2^j\rrbracket(\Delta_2^j), 0) \\ &= 0 \end{aligned}$$

It follows that $\mathbf{d}_{sb}^{\text{ls}}(s, t) \geq \llbracket\varphi\rrbracket(s) - \llbracket\varphi\rrbracket(t) > \epsilon$, which gives rise to a contradiction. \square

By combining the above three technical lemmas we obtain the following logical characterisation of the state-based bisimilarity metric.

Theorem 12. $\mathbf{d}_{sb} = \mathbf{d}_{sb}^{\text{ls}}$ and $K(\mathbf{d}_{sb}) = \mathbf{d}_{sb}^{\text{ld}}$ \square

Remark 13. In the proof of Lemma 11 we have constructed the formula

$$\varphi = \langle a \rangle \bigwedge_i (\psi_2^i \ominus \llbracket\psi_2^i\rrbracket(\Delta_2^i)) \quad (5)$$

by making use of conjunction and minus connectives for distribution formulae. This happens because in the presence of non-determinism state t may perform action a and then evolves into one of successor distributions Δ_2^i . If we confine ourselves to deterministic pLTSs, then state t will have a unique successor distribution Δ_2^i and therefore (5) can be simplified as $\varphi = \langle a \rangle \psi_2^i$. In this case, there is no need of conjunction and minus connectives for distribution formulae. Furthermore, if we fold $\llbracket\varphi\rrbracket$ into state formulae in Definition 5, distribution formulae can be completely dropped. In other words, for deterministic pLTSs, the state-based bisimilarity metric can be characterised by the following metric logic

$$\varphi ::= \top \mid \neg\varphi \mid \varphi \ominus p \mid \varphi_1 \wedge \varphi_2 \mid \langle a \rangle \varphi \quad (6)$$

Therefore, for deterministic pLTSs, the two-sorted logic in Definition 5 degenerates into the logic considered in [DGJP04, vBW05], as expected.

In [DAMRS08, CDAMR10] a bisimulation metric for game structures is characterised by a quantitative μ -calculus where formulae are valuated also on states and no distribution formula is needed. This is not surprising because games are deterministic: at any state s , if two players have chosen their moves, say a_1 and a_2 , then there is a unique distribution $\delta(s, a_1, a_2)$ to determine the probabilities of arriving at a set of destination states.

4 Distribution-Based Bisimulation Metric

The bisimilarity metric given in Definition 2 measures the distance between two states. Alternatively, it is possible to directly define a metric that measures subdistributions. In order to do so, we first define a transition relation between subdistributions.

Definition 14. We write $\Delta \xrightarrow{a} \Delta'$ if $\Delta' = \sum_{s \in [\Delta]} \Delta(s) \cdot \Delta_s$, where Δ_s is determined as follows:

- either $s \xrightarrow{a} \Delta_s$
- or there is no Θ with $s \xrightarrow{a} \Theta$, and in this case we set $\Delta_s = \varepsilon$.

Note that if $\Delta \xrightarrow{a} \Delta'$ then some (not necessarily all) states in the support of Δ can perform action a . For example, consider the two states s_2 and s_3 in Figure 1. Since $s_2 \xrightarrow{c} \overline{s_4}$ and s_3 cannot perform action c , the distribution $\Delta = \frac{1}{2}\overline{s_2} + \frac{1}{2}\overline{s_3}$ can make the transition $\Delta \xrightarrow{c} \frac{1}{2}\overline{s_4}$ to reach the subdistribution $\frac{1}{2}\overline{s_4}$.

Definition 15. A 1-bounded pseudometric d on $\mathcal{D}(S)$ is a distribution-based bisimulation metric if $||\Delta_1| - |\Delta_2|| \leq d(\Delta_1, \Delta_2)$ and for all $\Delta_1, \Delta_2 \in \mathcal{D}(S)$ and $\epsilon \in [0, 1)$ with $d(\Delta_1, \Delta_2) \leq \epsilon$, if $\Delta_1 \xrightarrow{a} \Delta'_1$ then there exists an $\Delta_2 \xrightarrow{a} \Delta'_2$ with $d(\Delta'_1, \Delta'_2) \leq \epsilon$.

The smallest (wrt. \sqsubseteq) distribution-based bisimulation metric, notation \mathbf{d}_{db} , is called *distribution-based bisimilarity metric*. Distribution-based bisimilarity [DFD15] is the kernel of the distribution-based bisimilarity metric.

It is not difficult to see that \mathbf{d}_{sb} is different from \mathbf{d}_{db} , as witnessed by the following example.

Example 16. Consider the states in Figure 1. We first observe that $\mathbf{d}_{db}(\overline{s_2}, \overline{t_3}) = 0$ because s_2 and t_3 can match each other's action exactly. Similarly, we have $\mathbf{d}_{db}(\overline{s_3}, \overline{t_4}) = 0$. Then it is easy to see that $\mathbf{d}_{db}(\frac{1}{2}\overline{s_2} + \frac{1}{2}\overline{s_3}, \frac{1}{2}\overline{t_3} + \frac{1}{2}\overline{t_4}) = 0$. Since $s_1 \xrightarrow{b} \frac{1}{2}\overline{s_2} + \frac{1}{2}\overline{s_3}$ and $\frac{1}{2}\overline{t_1} + \frac{1}{2}\overline{t_2}$, we infer that $\mathbf{d}_{db}(\overline{s_1}, \frac{1}{2}\overline{t_1} + \frac{1}{2}\overline{t_2}) = 0$. It, in turn, implies $\mathbf{d}_{db}(\overline{s}, \overline{t}) = 0$. We have already seen in Example 3 that $\mathbf{d}_{sb}(s, t) = \frac{1}{2}$. Therefore, the two distance functions \mathbf{d}_{sb} and \mathbf{d}_{db} are indeed different. \square

The rest of this section is devoted to a logical characterisation of \mathbf{d}_{db} . Consider the metric logic \mathcal{L}^{D*} whose formulae are defined below

$$\psi ::= \top \mid \neg\psi \mid \psi \ominus p \mid \psi_1 \wedge \psi_2 \mid \langle a \rangle \psi \quad (7)$$

This logic is the same as that defined in (6) except that now we only have distribution formulae. We will show that this logic can capture the distribution-based bisimilarity metric.

Definition 17. A formula $\psi \in \mathcal{L}^{D*}$ evaluates in $\Delta \in \mathcal{D}(S)$ as follows:

$$\llbracket \top \rrbracket(\Delta) = |\Delta|,$$

$$\begin{aligned}
\llbracket \neg\psi \rrbracket(\Delta) &= 1 - \llbracket \psi \rrbracket(\psi), \\
\llbracket \psi \ominus p \rrbracket(\Delta) &= \max(\llbracket \psi \rrbracket(\psi) - p, 0), \\
\llbracket \psi_1 \wedge \psi_2 \rrbracket(\Delta) &= \min(\llbracket \psi_1 \rrbracket(\Delta), \llbracket \psi_2 \rrbracket(\Delta)), \\
\llbracket \langle a \rangle \psi \rrbracket(\Delta) &= \max_{\Delta \xrightarrow{a} \Delta'} \llbracket \psi \rrbracket(\Delta'),
\end{aligned}$$

This induces a natural logical metric $\mathbf{d}_{db}^{\text{ld}}$ over subdistributions defined by

$$\mathbf{d}_{db}^{\text{ld}}(\Delta, \Theta) = \sup_{\psi \in \mathcal{L}^D} |\llbracket \psi \rrbracket(\Delta) - \llbracket \psi \rrbracket(\Theta)|$$

It turns out that $\mathbf{d}_{db}^{\text{ld}}$ coincides with \mathbf{d}_{db} . We split the proof of this coincidence result into two parts, to show that one metric is dominated by the other and vice versa.

Lemma 18. $\mathbf{d}_{db}^{\text{ld}} \subseteq \mathbf{d}_{db}$

Proof. Similar to the proof of Lemma 9. We proceed by structural induction on formulae. For any two subdistributions $\Delta_1, \Delta_2 \in \mathcal{D}(S)$, we prove that

$$|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$$

for all $\psi \in \mathcal{L}^{D*}$.

We first analyze the structure of ψ .

- $\varphi \equiv \top$. Then it is trivial to see that $|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| = ||\Delta_1| - |\Delta_2|| \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$.
- $\psi \equiv \neg\psi'$. Then $|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| = |\llbracket \psi' \rrbracket(\Delta_2) - \llbracket \psi' \rrbracket(\Delta_1)| \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$ where the inequality holds by induction.
- $\psi \equiv \psi' \ominus p$. There are four subcases and we consider one of them. Suppose $\llbracket \psi' \rrbracket(\Delta_1) > p$ and $\llbracket \psi' \rrbracket(\Delta_2) \leq p$, then $|\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2)| = |\llbracket \psi' \rrbracket(\Delta_1) - p| \leq |\llbracket \psi' \rrbracket(\Delta_1) - \llbracket \psi' \rrbracket(\Delta_2)| \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$ by induction.
- $\psi \equiv \psi_1 \wedge \psi_2$. Without loss of generality we assume that $\llbracket \psi \rrbracket(\Delta_1) \geq \llbracket \psi \rrbracket(\Delta_2)$. There are two possibilities:
 - If $\llbracket \psi_1 \rrbracket(\Delta_2) \leq \llbracket \psi_2 \rrbracket(\Delta_2)$, then $\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \leq \llbracket \psi_1 \rrbracket(\Delta_1) - \llbracket \psi_1 \rrbracket(\Delta_2) \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$, where the last inequality holds by induction.
 - Symmetrically, if $\llbracket \psi_2 \rrbracket(\Delta_2) \leq \llbracket \psi_1 \rrbracket(\Delta_2)$, then $\llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \leq \llbracket \psi_2 \rrbracket(\Delta_1) - \llbracket \psi_2 \rrbracket(\Delta_2) \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$.
- $\psi \equiv \langle a \rangle \psi'$. Let Δ'_1 be the distribution such that $\Delta_1 \xrightarrow{a} \Delta'_1$ and $\llbracket \langle a \rangle \psi' \rrbracket(\Delta_1) = \llbracket \psi' \rrbracket(\Delta'_1)$. Since \mathbf{d}_{db} is a distribution-based bisimulation metric, by definition there exists some Δ'_2 such that $\Delta_2 \xrightarrow{a} \Delta'_2$ and $\mathbf{d}_{db}(\Delta'_1, \Delta'_2) \leq \mathbf{d}_{db}(\Delta_1, \Delta_2)$. Without loss of generality we assume that $\llbracket \psi \rrbracket(\Delta_1) \geq \llbracket \psi \rrbracket(\Delta_2)$. It follows that

$$\begin{aligned}
& \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) \\
&= \llbracket \psi' \rrbracket(\Delta'_1) - \max_{\Delta_2 \xrightarrow{a} \Delta'_2} \llbracket \psi' \rrbracket(\Delta'_2) \\
&\leq \llbracket \psi' \rrbracket(\Delta'_1) - \llbracket \psi' \rrbracket(\Delta'_2) \\
&\leq \mathbf{d}_{db}(\Delta'_1, \Delta'_2) \quad \text{by induction on } \psi' \\
&\leq \mathbf{d}_{db}(\Delta_1, \Delta_2)
\end{aligned}$$

□

Lemma 19. $\mathbf{d}_{db} \sqsubseteq \mathbf{d}_{db}^{\text{ld}}$

Proof. Similar to the proof of Lemma 11. We show that $\mathbf{d}_{db}^{\text{ld}}$ is a distribution-based bisimulation metric. Let Δ_1, Δ_2 be any two subdistributions in $\mathcal{D}(S)$ and ϵ be any real number in the interval $[0, 1)$ with $\mathbf{d}_{db}^{\text{ld}}(\Delta_1, \Delta_2) \leq \epsilon$. Assume that $\Delta_1 \xrightarrow{a} \Delta'_1$ is an arbitrarily chosen transition from Δ_1 . We need to show that there exists some transition $\Delta_2 \xrightarrow{a} \Delta'_2$ with $\mathbf{d}_{db}^{\text{ld}}(\Delta'_1, \Delta'_2) \leq \epsilon$. Suppose for a contradiction that no a -transition from Δ_2 satisfies this condition. In other words, for each Δ_2^i with $\Delta_2 \xrightarrow{a} \Delta_2^i$ we have $\mathbf{d}_{db}^{\text{ld}}(\Delta'_1, \Delta_2^i) > \epsilon$. Then there must exist some formula $\psi_2^i \in \mathcal{L}^D$ such that $|\llbracket \psi_2^i \rrbracket(\Delta'_1) - \llbracket \psi_2^i \rrbracket(\Delta_2^i)| > \epsilon$. Furthermore, we can strengthen this condition to the following one

$$\llbracket \psi_2^i \rrbracket(\Delta'_1) - \llbracket \psi_2^i \rrbracket(\Delta_2^i) > \epsilon \quad (8)$$

because we can take the formula $\neg \psi_2^i$ in place of ψ_2^i in the case that $\llbracket \psi_2^i \rrbracket(\Delta'_1) < \llbracket \psi_2^i \rrbracket(\Delta_2^i)$. Let $\varphi = \langle a \rangle \bigwedge_i (\psi_2^i \ominus \llbracket \psi_2^i \rrbracket(\Delta_2^i))$. We infer that

$$\begin{aligned} \llbracket \varphi \rrbracket(\Delta_1) &= \max_{\Delta_1 \xrightarrow{a} \Delta} \llbracket \bigwedge_i \psi_2^i \ominus \llbracket \psi_2^i \rrbracket(\Delta_2^i) \rrbracket(\Delta) \\ &\geq \llbracket \bigwedge_i (\psi_2^i \ominus \llbracket \psi_2^i \rrbracket(\Delta_2^i)) \rrbracket(\Delta'_1) \\ &= \min_i \llbracket \psi_2^i \ominus \llbracket \psi_2^i \rrbracket(\Delta_2^i) \rrbracket(\Delta'_1) \\ &= \llbracket \psi_2^k \ominus \llbracket \psi_2^k \rrbracket(\Delta_2^k) \rrbracket(\Delta'_1) \quad \text{for some } k \\ &= \max(\llbracket \psi_2^k \rrbracket(\Delta'_1) - \llbracket \psi_2^k \rrbracket(\Delta_2^k), 0) \\ &> \epsilon \quad \text{by (8)} \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \llbracket \psi \rrbracket(\Delta_2) &= \max_{\Delta_2 \xrightarrow{a} \Delta_2^i} \llbracket \bigwedge_j (\psi_2^j \ominus \llbracket \psi_2^j \rrbracket(\Delta_2^j)) \rrbracket(\Delta_2^i) \\ &= \max_{\Delta_2 \xrightarrow{a} \Delta_2^i} \min_j \llbracket \psi_2^j \ominus \llbracket \psi_2^j \rrbracket(\Delta_2^j) \rrbracket(\Delta_2^i) \\ &= \max_{\Delta_2 \xrightarrow{a} \Delta_2^i} \min_j \max((\llbracket \psi_2^j \rrbracket(\Delta_2^i) - \llbracket \psi_2^j \rrbracket(\Delta_2^j)), 0) \\ &= 0 \end{aligned}$$

It follows that $\mathbf{d}_{db}^{\text{ld}}(\Delta_1, \Delta_2) \geq \llbracket \psi \rrbracket(\Delta_1) - \llbracket \psi \rrbracket(\Delta_2) > \epsilon$, which gives rise to a contradiction. □

By combining the previous two lemmas, we arrive at the following logical characterisation of the distribution-based bisimulation metric.

Theorem 20. $\mathbf{d}_{db} = \mathbf{d}_{db}^{\text{ld}}$

□

5 Related work

Metrics for probabilistic transition systems are first suggested by Giacalone *et al.* to formalize a notion of distance between processes. They are used also in [KN96, Nor97] to give denotational semantics for reactive models. De Vink and

Rutten [dVR99] show that discrete probabilistic transition systems can be viewed as coalgebras. They consider the category of complete ultrametric spaces. Similar ultrametric spaces are considered by den Hartog in [dH02].

Metrics for deterministic systems are extensively studied. Desharnais *et al.* [DGJP04] propose a logical pseudometric for labelled Markov chains, which is a reactive model of probabilistic systems. A similar pseudometric is defined by van Breugel and Worrell [vBW01b] via the terminal coalgebra of a functor based on a metric on the space of Borel probability measures. The metric of [DGJP04, vBW05] works for continuous probabilistic transition systems, while in this work we concentrate on discrete systems. Interestingly, van Breugel and Worrell [vBW01a] also present a polynomial-time algorithm to approximate their coalgebraic distances. Furthermore, van Breugel *et al.* propose an algorithm to approximate a behavioural pseudometric without discount [vBSW08]. In [FPP11] a sampling algorithm for calculating bisimulation distances in Markov decision processes is shown to have good performance. In [DAMRS07, DAMRS08] the probabilistic bisimulation metric on game structures is characterised by a quantitative μ -calculus. Algorithms for game metrics are proposed in [CDAMR10, Ram10].

Metrics for nondeterministic probabilistic systems are considered in [DJGP02], where Desharnais *et al.* deal with labelled concurrent Markov chains (similar to pLTSs, this model can be captured by the simple probabilistic automata of [Seg95]). They show that the greatest fixed point of a monotonic function on pseudometrics corresponds to the weak probabilistic bisimilarity of [PLS00].

In [SDC07] a notion of trace metric is proposed for pLTSs and a tool is developed to compute the trace metric. In [DCPP06] a notion of bisimulation metric is proposed that extends the approach of [DJGP02, DGJP04] to a more general framework called action-labelled quantitative transition systems.

In [DAFS09] de Alfaro *et al.* consider metric transition systems in which the propositions at each state are interpreted as elements of metric spaces. In that setting, trace equivalence and bisimulation give rise to linear and branching distances that can be characterised by quantitative versions of linear-time temporal logic [MP91] and μ -calculus [Koz83].

In [Yin02] Ying proposes a notion of bisimulation index for the usual labelled transition systems, by using ultrametrics on actions instead of using pseudometrics on states. He applies bisimulation indexes on timed CCS and real time ACP. But the deeper connection between [Yin02] and our work warrants some further studies.

In [GLT15] a notion of uniform continuity is proposed to be an appropriate property of probabilistic processes for compositional reasoning with respect to bisimulation metric semantics.

6 Concluding remarks

We have considered two behavioural pseudometrics for probabilistic labelled transition systems where nondeterminism and probabilities co-exist. They correspond to state-based and distribution-based bisimulations. Our modal charac-

terisation of the state-based bisimulation metric is much simpler than an earlier proposal by Desharnais *et al.* since we only use two non-expansive operators, negation and testing, rather than the general class of non-expansive operators. Our modal characterisation of the distribution-based bisimulation metric is new. The characterisations are shown to be sound and complete.

In the current work we have not distinguished internal actions from external ones. But it is not difficult to make the distinction and abstract away internal actions so as to introduce weak versions of bisimulation metrics. In a finite-state and finitely branching pLTS, the subdistributions reachable from a state by weak transitions is infinite but can be represented by the convex closure of a finite set [Den15]. This entails that the logical characterisation of weak bisimulation metrics would be similar to those presented here.

References

- CDAMR10. Krishnendu Chatterjee, Luca De Alfaro, Rupak Majumdar, and Vishwanath Raman. Algorithms for game metrics. *Logical Methods in Computer Science*, 6(3), 2010.
- DAFS09. Luca De Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Transactions on Software Engineering*, 35(2):258–273, 2009.
- DAMRS07. L. De Alfaro, R. Majumdar, V. Raman, and M. Stoelinga. Game relations and metrics. In *Proc. LICS’07*, pages 99–108. IEEE, 2007.
- DAMRS08. Luca De Alfaro, Rupak Majumdar, Vishwanath Raman, and Mariëlle Stoelinga. Game refinement relations and metrics. *arXiv preprint arXiv:0806.4956*, 2008.
- DCPP06. Yuxin Deng, Tom Chothia, Catuscia Palamidessi, and Jun Pang. Metrics for action-labelled quantitative transition systems. *ENTCS*, 153(2):79–96, 2006.
- Den15. Yuxin Deng. *Semantics of Probabilistic Processes: An Operational Approach*. Springer, 2015.
- DFD15. Yuxin Deng, Yuan Feng, and Ugo Dal Lago. On coinduction and quantum lambda calculi. In *Proceedings of the 26th International Conference on Concurrency Theory*, pages 427–440. LIPICS, 2015.
- DGJP04. Jose Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled markov processes. *Theor. Comput. Sci.*, 318(3):323–354, 2004.
- dH02. J. I. den Hartog. *Probabilistic Extensions of Semantical Models*. PhD thesis, Free University Amsterdam, 2002.
- DJGP02. Jose Desharnais, Radha Jagadeesan, Vineet Gupta, and Prakash Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proc. LICS’02*, pages 413–422. IEEE, 2002.
- DTW10. Pedro R. D’Argenio, Pedro Sánchez Terraf, and Nicolás Wolovick. Bisimulations for nondeterministic labeled markov processes. *CoRR*, abs/1011.3362, 2010.
- dVR99. E. P. de Vink and J. J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. *Theoretical Computer Science*, 221(1/2):271–293, 1999.

- FPP11. Norm Ferns, Prakash Panangaden, and Doina Precup. Bisimulation metrics for continuous markov decision processes. *SIAM J. Comput.*, 40(6):1662–1714, 2011.
- FZ14. Yuan Feng and Lijun Zhang. When equivalence and bisimulation join forces in probabilistic automata. In *Proceedings of the 19th International Symposium on Formal Methods*, volume 8442 of *Lecture Notes in Computer Science*, pages 247–262. Springer, 2014.
- GLT15. Daniel Gebler, Kim Guldstrand Larsen, and Simone Tini. Compositional metric reasoning with probabilistic process calculi. In *Proceedings of the 18th International Conference on Foundations of Software Science and Computation Structures*, volume 9034 of *Lecture Notes in Computer Science*, pages 230–245. Springer, 2015.
- Hen12. Matthew Hennessy. Exploring probabilistic bisimulations, part i. *Formal Aspects of Computing*, 24(4-6):749–768, 2012.
- HKK14. Holger Hermanns, Jan Krcál, and Jan Kretínský. Probabilistic bisimulation: Naturally on distributions. In *Proceedings of the 25th International Conference on Concurrency Theory*, volume 8704 of *Lecture Notes in Computer Science*, pages 249–265. Springer, 2014.
- HM85. Matthew Hennessy and Robin Milner. Algebraic laws for nondeterminism and concurrency. *J. ACM*, 32:137–161, 1985.
- KN96. M. Z. Kwiatkowska and G. J. Norman. Probabilistic metric semantics for a simple language with recursion. In *Proceedings of 21st International Symposium on Mathematical Foundations of Computer Science*, volume 1113 of *Lecture Notes in Computer Science*, pages 419–430. Springer-Verlag, 1996.
- Koz83. D. Kozen. Results on the propositional mu-calculus. *Theoretical Computer Science*, 27:333–354, 1983.
- KR58. L.V. Kantorovich and G.S. Rubinshtein. On a space of totally additive functions. *Vestn Len. Univ.*, 13(7):52–59, 1958.
- LS91. Kim G. Larsen and Arne Skou. Bisimulation through probabilistic testing. *Inf. Comput.*, 94:1–28, 1991.
- MP91. Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, 1991.
- Nor97. G. J. Norman. *Metric Semantics for Reactive Probabilistic Systems*. PhD thesis, University of Birmingham, 1997.
- PLS00. Anna Philippou, Insup Lee, and Oleg Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes in Computer Science*, pages 334–349. Springer-Verlag, 2000.
- Ram10. Vishwanath Raman. *Game Relations, Metrics and Refinements*. PhD thesis, University of California, 2010.
- SDC07. Lin Song, Yuxin Deng, and Xiaojuan Cai. Towards automatic measurement of probabilistic processes. In *Proc. QSI’07*, pages 50–59. IEEE, 2007.
- Seg95. Roberto Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, MIT, 1995.
- vBSW08. Franck van Breugel, Babita Sharma, and James Worrell. Approximating a behavioural pseudometric without discount for probabilistic systems. *Logical Methods in Computer Science*, 4(2), 2008.

- vBW01a. Franck van Breugel and James Worrell. An algorithm for quantitative verification of probabilistic transition systems. In *Proc. CONCUR'01*, volume 2154 of *LNCS*, pages 336–350. Springer, 2001.
- vBW01b. Franck van Breugel and James Worrell. Towards quantitative verification of probabilistic transition systems. In Fernando Orejas, Paul Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming*, volume 2076 of *LNCS*, pages 421–432. Springer, 2001.
- vBW05. Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331(1):115–142, 2005.
- Yin02. M. Ying. Bisimulation indexes and their applications. *Theoretical Computer Science*, 275(1/2):1–68, 2002.

A Convex Bisimulation Metric

For $\Pi \subseteq \mathcal{D}(S)$ we denote by $cc(\Pi)$ the convex closure of Π . If $\Delta \in cc(\text{der}(s, a))$ then we say Δ is a combined transition of s labelled by a , written as $s \xrightarrow{a}_c \Delta$.

Definition 21 (Convex bisimulation metric). *A 1-bounded metric d on S is a convex bisimulation metric if for all $s, t \in S$ and $\epsilon \in [0, 1)$ with $d(s, t) \leq \epsilon$, if $s \xrightarrow{a}_c \Delta$ then there exists a $\Delta' \in cc(\text{der}(t, a))$ such that $K(d)(\Delta, \Delta') \leq \epsilon$.*

The smallest (wrt. \sqsubseteq) convex bisimulation metric, denoted by \mathbf{d}_{cb} , is called *convex bisimilarity metric*. Convex bisimilarity equivalence (also called probabilistic bisimilarity) [Seg95] is the kernel of the convex bisimilarity metric.

Let us define the functor $F_{cb}: [0, 1]^{S \times S} \rightarrow [0, 1]^{S \times S}$ for $d: S \times S \rightarrow [0, 1]$ and $s, t \in S$ by

$$F_{cb}(d)(s, t) = \sup_{a \in A} \{H(K(d))(cc(\text{der}(s, a)), cc(\text{der}(t, a)))\}.$$

It can be shown that F_{cb} is monotone and its least fixed point is the convex bisimilarity metric.

Given a pLTS, we can saturate it with all possible combined transitions. In the saturated pLTS, convex bisimulation metric coincides with state-based bisimulation metric because of the following lemma.

Lemma 22. *d is a convex bisimulation metric if and only if for all $s, t \in S$ and $\epsilon \in [0, 1)$ with $d(s, t) \leq \epsilon$, if $s \xrightarrow{a}_c \Delta$ then there exists some $t \xrightarrow{a}_c \Delta'$ such that $K(d)(\Delta, \Delta') \leq \epsilon$.*

Proof. The “if” direction is trivial. The “only if” direction can be shown by making use of the following inequality:

$$K(d)\left(\sum_{i \in I} p_i \cdot \Delta_i, \sum_{i \in I} p_i \cdot \Delta'_i\right) \leq \sum_{i \in I} p_i \cdot K(d)(\Delta_i, \Delta'_i)$$

for any pseudometric d . This holds because

$$\begin{aligned} & K(d)\left(\sum_{i \in I} p_i \cdot \Delta_i, \sum_{i \in I} p_i \cdot \Delta'_i\right) \\ &= \max\left\{\sum_{u \in S} \left(\left(\sum_{i \in I} p_i \cdot \Delta_i\right)(u) - \left(\sum_{i \in I} p_i \cdot \Delta'_i\right)(u)\right)x_u \mid \right. \\ & \quad \left. x_u, x_{u'} \in [0, 1] \wedge x_u - x_{u'} \leq d(u, u')\right\} \\ &= \max\left\{\sum_{u \in S} \sum_{i \in I} p_i \cdot (\Delta_i(u) - \Delta'_i(u))x_u \mid x_u, x_{u'} \in [0, 1] \wedge x_u - x_{u'} \leq d(u, u')\right\} \\ &= \max\left\{\sum_{i \in I} p_i \sum_{u \in S} (\Delta_i(u) - \Delta'_i(u))x_u \mid x_u, x_{u'} \in [0, 1] \wedge x_u - x_{u'} \leq d(u, u')\right\} \\ &\leq \sum_{i \in I} p_i \cdot \max\left\{\sum_{u \in S} (\Delta_i(u) - \Delta'_i(u))x_u \mid x_u, x_{u'} \in [0, 1] \wedge x_u - x_{u'} \leq d(u, u')\right\} \\ &= \sum_{i \in I} p_i \cdot K(d)(\Delta_i, \Delta'_i) \end{aligned}$$

□

Consequently, a simple logical characterisation of convex bisimulation metric can be obtained by extending the logic in Section ?? with infinitary conjunctions and by interpreting the diamond modality with combined transitions, i.e. $\llbracket \langle a \rangle \psi \rrbracket(s) = \max_{s \xrightarrow{a}_c \Delta} \llbracket \psi \rrbracket(\Delta)$. Write \mathbf{d}_{cb}^1 for the metric induced by this extended logic.

Theorem 23. $\mathbf{d}_{cb} = \mathbf{d}_{cb}^1$

□

B Trace Metric

In this section we present a notion of trace metric that enjoys a straightforward logical characterisation.

B.1 Trace Metric

A trace tr is a string in the set A^* . We write ϵ for the empty trace, and $tr_1 \cdot tr_2$ for the concatenation of two traces tr_1 and tr_2 . Given a finitely branching pLTS, the maximum probability that state s can perform trace tr is defined as follows.

$$Pr(s, tr) = \begin{cases} 1 & \text{if } tr = \epsilon \\ \max_{s \xrightarrow{a} \Delta} \sum_{t \in S} \Delta(t) \cdot Pr(t, tr') & \text{if } tr = a \cdot tr' \end{cases}$$

Definition 24. For any two states $s, t \in S$, the trace distance between them is given as follows:

$$\mathbf{d}_t(s, t) = \sup_{tr \in A^*} |Pr(s, tr) - Pr(t, tr)|$$

Intuitively, the trace distance between s and t is the maximal difference between the maximum probabilities given by s and t to a same trace.

B.2 Logical Characterizations

Let \mathcal{L}^T be the set of formulae produced by the following grammar.

$$\varphi ::= \top \mid \langle a \rangle \varphi$$

A formula evaluates in a state s as follows:

$$\begin{aligned} \llbracket \top \rrbracket(s) &= 1 \\ \llbracket \langle a \rangle \varphi \rrbracket(s) &= \max_{s \xrightarrow{a} \Delta} \sum_{t \in S} \Delta(t) \cdot \llbracket \varphi \rrbracket(t) \end{aligned}$$

This logic induces a natural metric \mathbf{d}_t^1 .

$$\mathbf{d}_t^1(s, t) = \sup_{\varphi \in \mathcal{O}_{tr}} |\llbracket \varphi \rrbracket(s) - \llbracket \varphi \rrbracket(t)|$$

Theorem 25. $\mathbf{d}_t = \mathbf{d}_t^1$

Proof. It is easy to see that there is a bijection f between A^* and \mathcal{L}^T .

$$\begin{aligned} f(\epsilon) &= \top \\ f(a \cdot tr) &= \langle a \rangle f(tr) \end{aligned}$$

For any state s and trace tr , we can show that $Pr(s, tr) = \llbracket f(tr) \rrbracket(s)$ by induction on the length of tr . The required result now easily follows. \square